

# White Paper - Pitfalls when enriching a knowledge base.

All business and government departments are facing increasing pressure to make decisions faster and to respond to issues immediately. Due to these demands, organisations are turning to Operational Intelligence to help solve these problems, however, what are the pitfalls along the way and how do you avoid them?

## Fujitsu's Operational Intelligence Team

Fujitsu Australia's operational intelligence team is leading the way in delivering operational intelligence solutions within Federal and State government agencies. Fujitsu has partnered with a number of major vendors in order to deliver customer requirements, including IBM i2 and Palantir Technologies. Whether it is the initial design, implementation, support or user training, Fujitsu is able to provide you a solution and avoid pitfalls.

### Introduction

A knowledge based (KB) system consists of information that represents facts about the world and domain of interest that can be used to deduce new facts. To achieve this, knowledge should be enriched and therein lays the problem, how do we best go about enriching this knowledge? If your organisation has a KB or has intentions to invest in a KB, think twice before you make design and workflow decisions as poorly thought through decisions can lead to the detriment of the KB.

We have seen a common trend in the industry in which the KB transforms from a rich and performant system into to a polluted and sluggish data store, which could have been avoided with smarter design decisions. This paper takes a deeper look into two practical challenges faced by organisations when enriching the KB and then provides recommendations to solve these issues.

### *Pitfall 1 - Resist the desire to store everything*

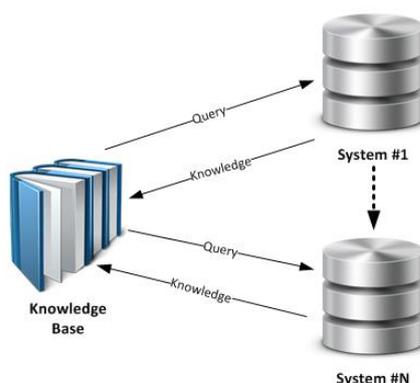
#### The Challenge

Storing enriched knowledge on a single platform, which is also your primary knowledge base, has its challenges. Due to poor decision making, organisations have ended up with a KB that contains copies of databases from external systems, all in the name of "knowledge enrichment". You might have all your data in one database, but the decision to synchronise information from another system into the KB can create a range of complications.

#### Background

Synching data from one system to the KB causes data to be out of date, which does not allow real time analysis and is a fundamental requirement of Operational Intelligence (OI) systems. Having data in one place has its advantages, such as less overhead in terms of authorisation checks, auditing is simpler, not needing to query external systems and generating basic reports might be easier. However, if your goal is to have immediate access to data, it can be a real issue.

Additionally, if you're of the view that a KB should know as much as it can "just in case", then this can lead to problems for your OI systems



**Figure 1 - Bulk importing**

because it can pollute existing knowledge, create data duplication and make finding useful data harder for the analyst.

Figure 1 describes how data from different systems are imported into the KB, ill-conceived to be knowledge because it actually has more characteristics of data loading than on providing real time analysis capability.

By effectively having copies of databases, you introduce inconsistency issues and have to develop new workflows to keep this new data in sync. With large volumes of data, even the best of systems available today will struggle to keep data in sync in real time during business hours. The typical work around is to then "refresh" this data nightly after hours. This job can get slower as more data is added to the KB or the source system.

Now, putting yourself in the analyst's shoes, when you query the KB and get results, a thorough analyst would wonder "is this the latest?". With doubts in their mind, the analyst is likely to query the source system directly to double check. Though this was not the intention, redundant steps have been introduced and extra load on systems that was otherwise not required. This could increase hardware and software licences, which further increases the total cost of the system.

Consider the example of having a KB about country information. You could store the countries name and rough population count in the KB but you would not want to necessarily store the average temperature of each city per day. Instead, you would query the current temperature from the external system directly, which will give you the most accurate result. If you did need the average temperature, then again it makes more sense to query the external system instead and calculate the average on the fly, which will give you the most up to date answer.

The same philosophy of enriching data and storing it back into the KB rather than constantly performing "data loading" applies to this context as well.

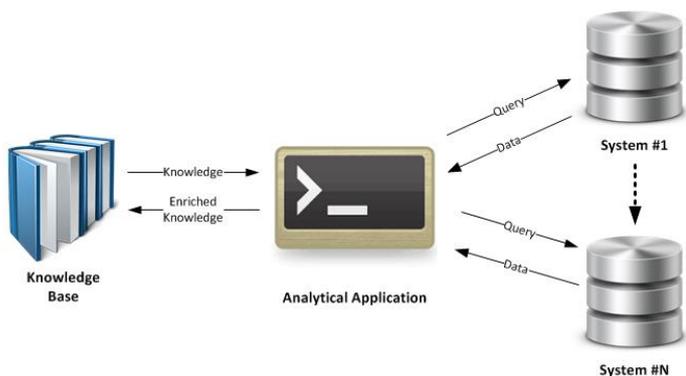
When importing large databases, if data is not mapped correctly it can be difficult to undo changes because of how KB's work (nothing really is deleted). Note that subtle mistakes are not always obvious and pop up months later, which require a lot of damage control for large KBs.

**Solution**

Software platforms available today do much more than 'store data' and their feature sets continue to grow. If the platform you are using has tools to query external systems and import them as knowledge then that is a useful capability. However, if this is not available or is limited and does not meet your requirement, you might be forced to consider building your own tools (and do so with caution).

Developing new tools introduces a whole range of issues like authorisation, discoverability, traceability, sustainment of the additional tools and implementing existing security models, which may not be an easy task due to complexity in these environments. However, if designed well, it can avoid data pollution, duplication and performance problems to name a few. We have seen customers achieve this successfully.

*Pitfall 2 - Using specialised tools to enrich knowledge*



**Figure 2 - Analytical Application**

**The Challenge**

Proceed with caution when developing new tools. If the tools provided by your knowledge platform do not meet your requirements, it is tempting to develop a program or script to solve the issue. Large organisations have hundreds of projects running in parallel so visibility of new tools being developed will easily get lost and add a lack of communication to the mix, you are likely to have more than one tool that does the same thing and teams re-inventing the wheel.

**Background**

Performing enrichment in a separate application will allow an analyst

to store truly enriched knowledge back into the KB. When the tool is developed to meet specific use cases, then you can appreciate that the tool could perform specialised crafted queries and allow the analyst to post process this data so there is only a precise piece of knowledge being promoted.

**Solution**

We know that first principles dictate to use the right tool for the job. To create an environment that does not have duplicate tools, strong governance and a clear vision across the organisation is critical. With enrichment tools in place, other systems are queried, the data validated, questioned, and promoted back as enriched knowledge. If you consider quality over quantity then you are more likely to avoid the pitfalls discussed in this paper.

**Operational Intelligence Services**

Fujitsu's Operational Intelligence Team is able to support our customers by using our experience in delivering systems Australia wide, and delivering the following services:

**Analysis and Design**

- Business Analysis
- Solution Architecture
- System Analysis
- Technical design
- Documentation
- Analyst training
- System testing

**Development**

- Palantir integration and development.
- IBM i2 integration and development.
- Java / J2EE
- Database design and development.
- Hadoop / NoSQL
- Bespoke development
- Web development

**System Support**

- On-going support of the production system.
- Out of hours support.

As well as, software licencing, consulting, programme and project management.

**Contact us**

For additional information, please contact [Operational.Intelligence@au.fujitsu.com](mailto:Operational.Intelligence@au.fujitsu.com)

**Contact**

Fujitsu Australia  
 Address: Level 3 , 4 National CCT Barton  
 ACT 2600 Australia  
 Phone: +61-2-6250 9600  
 E-mail:Operational.Intelligence@au.fujitsu.com  
 Website: www.fujitsu.com/au  
 2015-04-22

**About Fujitsu**

Fujitsu Australia and New Zealand is a leading service provider of business, information technology and communications solutions. As the third largest ICT company in the Australian and New Zealand marketplace, we partner with our customers to consult, design, build, operate and support business solutions. From strategic consulting to application and infrastructure solutions and services, Fujitsu Australia and New Zealand has earned a reputation as the single supplier of choice for leading corporate and government organisations. Fujitsu Australia Limited and Fujitsu New Zealand Limited are wholly owned subsidiaries of Fujitsu Limited (TSE: 6702).